

## Regional calibration of a watershed model

W. FERNANDEZ, R. M. VOGEL &  
A. SANKARASUBRAMANIAN

*Department of Civil and Environmental Engineering, Tufts University, Medford,  
Massachusetts 02155, USA*

e-mail: [rvogel@tufts.edu](mailto:rvogel@tufts.edu)

**Abstract** As watershed models become increasingly sophisticated and useful, there is a need to extend their applicability to locations where they cannot be calibrated or validated. A new methodology for the regionalization of a watershed model is introduced and evaluated. The approach involves calibration of a watershed model to many sites in a region, concurrently. Previous research that has sought to relate the parameters of monthly water balance models to physical drainage basin characteristics in a region has met with limited success. Previous studies have taken the two-step approach: (a) estimation of watershed model parameters at each site, followed by (b) attempts to relate model parameters to drainage basin characteristics. Instead of treating these two steps as independent, both steps are implemented concurrently. All watershed models in a region are calibrated simultaneously, with the dual objective of reproducing the behaviour of observed monthly streamflows and, additionally, to obtain good relationships between watershed model parameters and basin characteristics. The approach is evaluated using 33 basins in the southeastern region of the United States by comparing simulations using the regional models for three catchments which were not used to develop the regional regression equations. Although the regional calibration approach led to nearly perfect regional relationships between watershed model parameters and basin characteristics, these “improved” regional relationships did not result in improvements in the ability to model streamflow at ungauged sites. This experiment reveals that improvements in regional relationships between watershed model parameters and basin characteristics will not necessarily lead to improvements in the ability to calibrate a watershed model at an ungauged site.

### Calage régional d'un modèle de bassin hydrologique

**Résumé** Les modèles de bassins hydrologiques devenant de plus en plus complexes et utiles, il est désormais nécessaire que leur application soit étendue à des régions où il est difficile de les caler et de les valider. Cet article présente et évalue une nouvelle méthode de régionalisation de modèle qui réalise un calage simultané sur de nombreux sites d'une même région. Les recherches antérieures, qui essayaient d'associer les éléments d'un bilan hydrologique mensuel à des caractéristiques physiques du bassin de drainage correspondant, n'ont connu qu'un succès limité. Ces recherches utilisaient une approche en deux temps: (a) les paramètres du modèle étaient estimés pour chaque site d'étude, puis (b) on tentait de relier ces paramètres aux caractéristiques physiques du bassin. Au lieu de réaliser ces deux étapes indépendamment l'une de l'autre, nous les traitons simultanément. Nous calons en parallèle tous les modèles de bassin de toute une région avec pour double objectif de reproduire le comportement du débit mensuel et d'obtenir une bonne relation entre les paramètres du bilan hydrologique et les caractéristiques des bassins. Cette approche a été validée sur 33 bassins du sud-est des Etats-Unis en étudiant les simulations utilisant l'approche régionale de trois bassins qui n'avaient pas été utilisés pour l'établissement des équations régionales de régression. Notre méthode montre une relation presque parfaite entre les paramètres du modèle et les caractéristiques du bassin. L'utilisation de ces relations régionales “améliorées” n'a pas cependant pas permis d'améliorer notre capacité de création d'un modèle permettant d'évaluer les débits aux sites pour lesquels nous ne disposons pas de mesures physiques. Notre recherche souligne donc combien il est difficile de donner un sens physique aux paramètres d'un bassin hydrologique issus d'un calage.

## INTRODUCTION

As watershed models, computer technology and hydroclimatological data evolve, there is an ever increasing need to apply models for watersheds where streamflow data are unavailable. Without streamflow data, a watershed model cannot be calibrated; hence regional methods are needed which relate easily measured watershed characteristics to watershed model parameters. Abdulla & Lettenmaier (1997), Sefton & Howarth (1998), Xu & Singh (1998), Seibert (1999), Post & Jakeman (1999), and Xu (1999) provide recent reviews of dozens of studies which have used regionalization methods to relate parameters of rainfall-runoff models to watershed characteristics. Previous regionalization studies have focused on a wide range of hydrological models ranging from complex hourly and daily watershed models to the more parsimonious monthly water balance models.

Although each previous study attempted to regionalize a different watershed model, all studies to date follow the same general approach. First a watershed model is calibrated to climate and streamflow data available at each site in the region. This is followed by the application of a regionalization method which attempts to relate optimized model parameters to watershed characteristics. The most common regionalization method to date is bivariate and multivariate regression, although other methods have been attempted. Tung *et al.* (1997) recommend use of multivariate statistical methods which can account for the correlation structure among the watershed model parameters. Vandewiele & Elias (1995) found that kriging led to an improvement over multivariate regression for estimation of parameters of monthly water balance models at ungauged sites. Burn & Boorman (1992) showed that use of a watershed clustering algorithm to quantify watershed similarity led to improvements over the use of multivariate regression for estimation of two parameters of a unit hydrograph at ungauged sites in the United Kingdom. Post *et al.* (1998) document that incorporation of a regional relationship between annual runoff and forest density into model calibrations led to significant improvements in model application at ungauged locations. Servat & Dezetter (1993) found that it was easier to relate watershed model parameters to landscape attributes for parsimonious watershed models than for watershed models which are overparameterized. Even when one attempts to regionalize a very reasonable and parsimonious watershed model, results are still mixed (Post & Jakeman, 1999).

So far all previous watershed model regionalization studies have met with limited success. Kuczera & Mroczkowski (1998) suggested that attempts to regionalize watershed model parameters for the purpose of application to ungauged catchments will be *virtually impossible* due to the existence of multiple optimal model parameter sets and a high degree of correlation among model parameters. As a consequence, there exist many possible model parameter sets which produce virtually indistinguishable simulated streamflow sequences. The main idea of this research is to choose among those virtually indistinguishable parameter sets so as to maximize the “goodness of fit” of regional relationships between model parameters and drainage basin characteristics. Instead of choosing parameters which minimize the model residuals alone, the goal of this study is to both minimize model residuals *and* maximize the goodness of fit of relationships between model parameters and basin characteristics, concurrently. Naturally, this approach is computationally intensive, because all sites in the region

must be calibrated concurrently; however, recent advances in computer technology and nonlinear optimization algorithms enable one to readily implement this approach.

The methodology introduced here could be applied to any watershed model. In fact, this methodology could also be applied to the regionalization of other hydrological models including flood frequency, low flow frequency, and stochastic streamflow models. This initial study focuses on the regionalization of a four-parameter monthly water balance model for a region made up of 33 sites in the southeastern United States. The following sections outline the water balance model, the regional hydroclimatological database, and the regional calibration methodology followed by validation experiments which provide an evaluation of the overall methodology.

## A MONTHLY WATER BALANCE MODEL

Alley (1984), Vandewiele *et al.* (1992), Vandewiele & Ni-Lar-Win (1998) and Xu & Singh (1998) compared the performance of numerous alternative monthly water balance models and concluded that a three to five parameter model is sufficient to reproduce most of the information in a hydrological record on a monthly scale in humid regions. In those comparisons, all monthly models performed credibly and none stood out as clearly superior. This study focuses on the “abcd” model introduced by Thomas (1981) and Thomas *et al.* (1983) because it is comparable with other water balance models and each of its parameters has a physical interpretation.

### The “abcd” model

The “abcd” model is a nonlinear watershed model which accepts precipitation and potential evapotranspiration as input, producing streamflow as output. Internally, the model also represents soil moisture storage, groundwater storage, direct runoff, groundwater outflow to the stream channel and actual evapotranspiration. It was originally introduced by Thomas (1981) and Thomas *et al.* (1983) as a suitable model structure for performing regional water resource assessment using an annual time scale. The “abcd” model was later compared with numerous monthly water balance models leading to its recommendation by Alley (1984, 1985). Vandewiele *et al.* (1992) also found that the “abcd” model compares favourably with several other more recent monthly water balance models. The “abcd” model is unrelated to, and has a completely different structure from the linear “abc” model introduced by Fiering (1967) for pedagogic purposes.

The “abcd” model defines two state variables:  $W_t$ , termed “available water” and  $Y_t$ , “evapotranspiration opportunity”. Available water is defined as:

$$W_t = P_t + S_{t-1} \quad (1)$$

where  $P_t$  is precipitation during period  $t$  and  $S_{t-1}$  is soil moisture storage at the beginning of period  $t$ . Evapotranspiration opportunity is water which will eventually leave the basin in the form of evapotranspiration and is defined as:

$$Y_t = E_t + S_t \quad (2)$$

where  $E_t$  represents actual evapotranspiration during period  $t$  and  $S_t$  represents soil moisture storage at the end of period  $t$ . Evapotranspiration opportunity,  $Y_t$  is postulated as a nonlinear function of “available water”  $W_t$  using:

$$Y_t(W)_t = \frac{W_t + b}{2a} - \sqrt{\left(\frac{W_t + b}{2a}\right)^2 - \frac{W_t b}{a}} \quad (3)$$

This function simply assures that  $Y_t \leq W_t$ ,  $dW(0)/dY = 1$  and  $dW(\infty)/dY = 0$ . In fact, the upper limit on  $W_t$  is  $b$ . Thomas *et al.* (1983) note that “beside these properties, the function  $Y(W)$  has no particular significance”.

Allocation of available water,  $Y_t$ , between  $E_t$  and  $S_t$  is accomplished by assuming that the rate of loss of soil moisture to evapotranspiration is proportional to the soil moisture storage, so that  $dS/dt = -PE \cdot S/b$ . Solving this differential equation and assuming  $S_{t-1} = Y_t$  leads to:

$$S_t = Y_t \exp(-PE_t / b) \quad (4)$$

The difference between available water and evapotranspiration opportunity,  $W_t - Y_t$ , is also the sum of groundwater recharge and direct runoff. The parameter  $c$  allocates the quantity  $W_t - Y_t$  between groundwater recharge  $c(W_t - Y_t)$  and direct runoff  $(1 - c)(W_t - Y_t)$ . Finally, groundwater discharge to the stream channel is modelled as  $dG_t$ , where  $d$  is the fourth model parameter and  $G_t$  is groundwater storage at the end of period  $t$ . Groundwater storage is modelled using the continuity equation. Groundwater storage at the end of period  $t$  is equal to previous storage, plus groundwater recharge, less groundwater outflow, so that:

$$G_t = G_{t-1} + c \cdot (W_t - Y_t) - dG_t \quad (5)$$

Finally, streamflow  $Q_t$  is the sum of direct runoff and groundwater discharge:

$$Q_t = (1 - c) \cdot (W_t - Y_t) + dG_t + \varepsilon_t \quad (6)$$

where  $\varepsilon_t$  represents model error in month  $t$ .

### Regional physical relationships for “abcd” model parameters

The primary goal of this study is to test a new method for the calibration and regionalization of watershed models. In this initial study, an attempt is made to use as much information as is available to relate watershed model parameters to basin characteristics. Normally, when one attempts to regionalize a watershed model for use at ungauged sites, one includes only landscape attributes which are easily measured from digital elevation maps, soil maps, climate atlases and other existing sources of information. This enables estimation of watershed model parameters at ungauged sites, where presumably no streamflow data are available. Since the main objective of this study is to develop a methodology for the regionalization of a watershed model, the need to develop usable relationships at ungauged sites is not considered as a goal. Subsequent research will concentrate on the development of

regional relationships which use basin descriptors, which are easily measured at an ungauged site. Instead, this study uses several basin descriptors which require an analysis of streamflow data.

The “abcd” model has four parameters  $a$ ,  $b$ ,  $c$  and  $d$ , each having some physical interpretation. The parameter  $a$  ( $0 \leq a \leq 1$ ) reflects the “propensity of runoff to occur before the soil is fully saturated” (Thomas *et al.*, 1983). As shown by Alley (1984), the parameter  $a$  was found to fall in the range [0.95, 0.99] across broad regions of the United States. Runoff is expected to decrease as soil permeability increases, hence the parameter  $a$  is modelled using the regional regression model:

$$a = \alpha_a - \beta_a \cdot P \quad (7)$$

where  $P$  is basin permeability and  $\alpha_a$  and  $\beta_a$  are regional regression model parameters. Values of  $P$  are obtained from a digital grid of soil characteristics developed for the conterminous United States by Wolock (1997).

The parameter  $b$  is an upper limit on the sum of actual evapotranspiration and soil moisture storage in a given month. Presumably this parameter depends on the ability of the catchment to hold water within the upper soil horizon. In this study,  $b$  is modelled using the physical relationship:

$$b = \alpha_b + \beta_b \cdot P \quad (8)$$

where  $P$  is basin permeability and  $\alpha_b$  and  $\beta_b$  are regional regression model parameters.

The parameter  $c$  is equal to the fraction of streamflow which arises from groundwater discharge in a given month. Over the long term:

$$c = \frac{E[dG_t]}{E[Q_t]} = BFI \quad (9)$$

where  $BFI$  is the baseflow index used commonly in studies which develop relationships between drainage basin characteristics and groundwater discharge to a stream channel (see for example Gustard *et al.*, 1992). An algorithm is employed which was developed by the Institute of Hydrology (1980) to estimate the average annual  $BFI$  from the same records of daily streamflow used to calibrate the monthly water balance models. This algorithm is not based on the theory of groundwater outflow, hence equation (9) is not expected to hold exactly. Instead  $c$  is modelled using:

$$c = \alpha_c + \beta_c \cdot BFI \quad (10)$$

where  $\alpha_c$  and  $\beta_c$  are regional regression model parameters. Other algorithms for estimation of  $BFI$  are available, ranging from the use of the digital filtering algorithm introduced by Nathan & McMahon (1990) to the more theoretically based algorithm introduced by Rutledge & Daniel (1994). Gustard *et al.* (1992) review numerous studies which document the value of the  $BFI$  in regional flood and low flow studies. Burn & Boorman (1993) also found the  $BFI$  useful for estimation of unit hydrograph parameters at partially gauged sites.

One can easily show that the reciprocal of the parameter  $d$  is equal to the average groundwater residence time. Vogel & Kroll (1992, 1996) and others have shown that during baseflow conditions, when direct runoff is negligible and when groundwater

outflow is linearly proportional to groundwater storage,  $Q_t = dG_t$ ; as is assumed in the “abcd” model (equation (5)), streamflow follows the simple recursion:

$$Q_t = K_b Q_{t-1} \quad (11)$$

where  $K_b$  is termed the baseflow recession constant. Vogel & Kroll (1996) and equations (10) and (12) show that under these conditions equation (6) reduces to  $Q_t = -\ln(K_b) \cdot G_t$ . The regional regression for  $d$  is:

$$d = \alpha_d - \beta_d \cdot \ln(K_b) \quad (12)$$

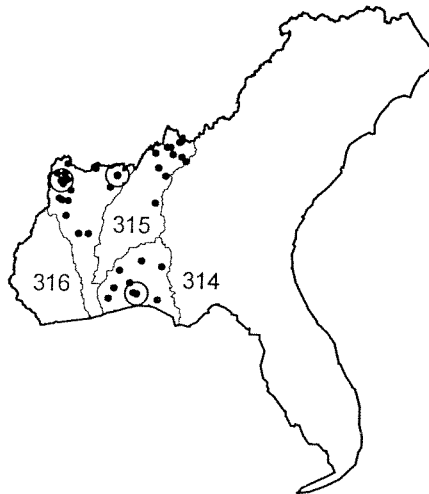
Estimates of the daily baseflow recession constant  $K_b$  are obtained for each of the rivers in this study using the estimator  $K_{b5}$  introduced by Vogel & Kroll (1996). This estimator assumes, as does the “abcd” model, that the groundwater aquifer acts like a linear reservoir. This estimator of  $K_b$  was favoured among several baseflow recession estimators compared by Vogel & Kroll (1996).

## DATA SOURCES

The “abcd” model requires time series of monthly precipitation, potential evapotranspiration and streamflow to enable calibration. The following sections describe these data sources.

### Monthly streamflow data

The streamflow data set consists of records of average monthly streamflow at 33 sites located in the southeastern region of the United States. Figure 1 uses solid and open circles to illustrate the location of the 30 calibration and three validation sites, respectively. Streamflow data were obtained from the hydroclimatologic data network



**Fig. 1** Location of the 33 US Geological Survey HCDN streamgauges located in Region 3 within Subregions 314, 315 and 316.

(HCDN), developed by the US Geological Survey (USGS) (Slack *et al.*, 1993). Streamflow gauges included in the HCDN are intended for use in climate sensitive studies and represent only a small subset of all streamflow data available in electronic form from the USGS. Screening procedures were used to assure that HCDN flow records are representative of “natural” or “unregulated” flow regimes. Streamflow records were only included which contained no overt adjustment of “natural” monthly flows by flow diversion, augmentation, groundwater pumping, or other forms of regulation and only measured discharge values are used, with no reconstructed or estimated flow records. All USGS streamflow records are organized into 18 regions and 220 subregions within the conterminous United States. All HCDN sites within subregions 314, 315 and 316 located within region 3, as illustrated in Fig. 1 were included.

The record lengths for the 33 stations ranged from 19 to 37 years with an average of 30.4 years. Drainage areas ranged from 155 to 39 847 km<sup>2</sup> with an average drainage area of 3031 km<sup>2</sup>. The average watershed elevation ranged from 60 m to 584 m with an average value of 207 m above mean sea level.

### Monthly climate data

The average annual precipitation for the 33 watersheds ranges from 1316 to 1640 mm with an average value of 1435 mm. Spatially weighted monthly time series of precipitation and potential evapotranspiration over the period 1951–1988 were developed using a geographic information system, a digital elevation map, and digital monthly time series grids for precipitation, minimum and maximum monthly temperature. The monthly precipitation, minimum monthly temperature and maximum monthly temperature time series were obtained from 0.5 degree digital time series grids using the PRISM (Daly *et al.*, 1994, 1997) climate analysis system. These grids were resampled to 0.1 degrees using bilinear interpolation. Spatially averaged values of each climate characteristic over each basin were obtained using the PRISM digital time series grids and watershed boundaries derived from a 1-km digital elevation map of the United States. The digital precipitation and temperature time series grids were generated using the PRISM modelling system (Daly *et al.*, 1994, 1997). PRISM Parameter-elevation Regressions on Independent Slopes Model is a climate analysis system that uses point data, a digital elevation model (DEM), and other spatial information to generate gridded estimates of annual, monthly, and event-based climatic parameters. It has been designed to accommodate difficult climate mapping situations in innovative ways. These include vertical extrapolation of climate well beyond the lowest or highest observation, reproducing gradients caused by rain shadows and coastal effects, and assessing the varying effects of terrain barriers on precipitation. Originally developed in 1991 for precipitation estimation, PRISM has been generalized and applied successfully to temperature, snowfall, growing degree-days, and weather generator parameters, among others (Taylor *et al.*, 1997).

### Monthly potential evapotranspiration

The spatially-averaged time series of monthly temperatures were combined with estimates of extraterrestrial solar radiation for each basin to obtain time series of

monthly potential evapotranspiration (PE) for each basin using the Hargreaves (Hargreaves & Samani, 1982) method. Extraterrestrial solar radiation was estimated for each basin by computing the solar radiation over  $0.1^\circ$  grids using a method introduced by Duffie & Beckman (1980), and then summing those estimates for each river basin. Even though it is only based on temperature and solar radiation measurements, numerous studies have shown that the Hargreaves method performs well when compared with other more complex methods. For example, the Hargreaves method was the highest ranked temperature based method for computing PE reported in the ASCE Manual 70 analysis (Jensen *et al.*, 1990). Allen (1993) showed that the Hargreaves method performs well in a wide range of latitudes and climates for periods of five days or longer, without significant error. Among all temperature-based methods, the Hargreaves method is the only one recommended by Shuttleworth (1993).

## MODEL CALIBRATIONS

Traditional approaches to model calibration assume that the primary objective is to obtain a “best fit” to the streamflows at each site; thus the objective function tends to focus on the model residuals at each site. The traditional calibration objective function treats each site independently even if the goal is to obtain a regional hydrological model. The idea of this study is to modify the objective function to reflect the fact that one’s interest is in both a “best fit” to the streamflows at each site and a “best fit” to the regional relationships which relate model parameters to watershed characteristics.

Two approaches to model calibration are compared: (a) traditional automatic calibration which estimates model parameters at each site which yield a “best fit” to streamflow observations, and (b) a regional calibration methodology which estimates model parameters at all sites concurrently in an effort to obtain a good fit to streamflows at all sites while simultaneously obtaining a good fit to the relationship between model parameters and watershed characteristics.

### At-site calibrations

Calibration algorithms have evolved considerably and it is now common practice to use a specially designed optimization algorithm, such as the shuffled complex evolution (SCE) algorithm developed at the University of Arizona for calibration of a watershed model (Duan *et al.*, 1992). Unfortunately, most algorithms such as the SCE algorithm are suited to calibration of a hydrological model at a single site and are not suited to the computational burdens posed by the regional calibration methodology introduced later in this paper. Instead a generalized reduced gradient nonlinear programming algorithm was employed: the Premium Solver Plus Version 3.5 (Frontline Systems, 1999) is an extension to the standard Microsoft Excel Solver, with the capacity to solve much larger problems, up to 1000 variables, at speeds anywhere from three to 100 times faster than the standard Solver. This algorithm was employed to calibrate the “abcd” model to the climate and streamflow traces at each of the 30 watersheds. This approach is termed the “at-site” calibration methodology. In this case, the objective at each site is to:



$$\text{Minimize } \sum_{t=1}^n \left( \ln(Q_t) - \ln(\hat{Q}_t) \right)^2 \quad (13)$$

where  $Q_t$  is observed monthly streamflow in month  $t$ ,  $\hat{Q}_t$  is modelled monthly streamflow in month  $t$  (from equation (6)) and  $n$  is the number of months of data available for calibration. The sum of the difference between logarithms of observed and modelled streamflow are minimized so as to give roughly equal weight to wet and dry months. Otherwise, without taking logarithms, reproduction of monthly mean flows during the dry summer months is poor.

At each site, the initial soil moisture storage  $S_0$  and the initial groundwater storage  $G_0$  are constrained to equal their average modelled values during the month of September, because model simulations always begin at the start of the water year on 1 October. Therefore  $G_0$  and  $S_0$  represent the average ending groundwater and soil moisture storage, respectively, in September. This approach is physically plausible and avoids the need to optimize two extra model parameters, instead treating them as model constraints.

### Regional calibrations

The traditional at-site approach described above, treats each site independently in an effort to obtain the best possible calibration at each site. The regional calibration approach attempts to get the best possible calibration at each site while simultaneously obtaining the best possible regional relationships between model parameters and basin characteristics. In this case the objective is to:

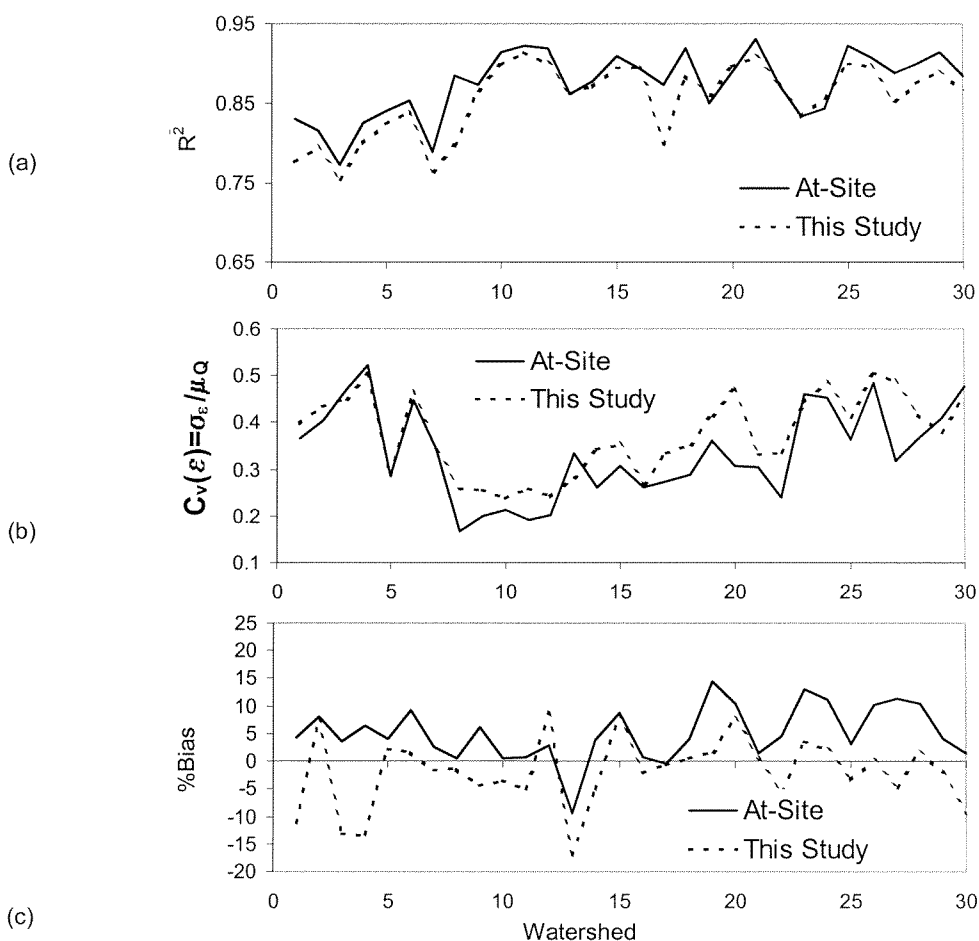
$$\text{Maximize } \left[ \frac{1}{m} \sum_{i=1}^m R_i^2 \right] + \left[ \frac{R_a^2 + R_b^2 + R_c^2 + R_d^2}{4} \right] \quad (14)$$

where there are  $m = 30$  sites in the region,  $R_i^2$  represents the coefficient of determination for site  $i$  which measures the goodness of fit of the logarithms of the modelled flows at site  $i$  and  $R_a^2$ ,  $R_b^2$ ,  $R_c^2$ , and  $R_d^2$  represent the coefficient of determination associated with each of the regression models for the model parameters  $a$ ,  $b$ ,  $c$  and  $d$  given in equations (7), (8), (10) and (12), respectively. The idea of the objective function in equation (14) is to maximize the average goodness of fit of the "abcd" model across all sites as well as to maximize the average goodness of fit of the four regional regression models. The coefficient of determination is employed as a measure of the goodness of fit, not because it is the best overall criterion, but because it provides a simple and equal weighting scheme for the two concurrent objectives.

To implement the regional calibration approach, one could use the SCE algorithm (Duan *et al.*, 1992); however, since there are now  $30(4) = 120$  model parameters to optimize, this approach is computationally infeasible. Instead a generalized reduced gradient nonlinear programming algorithm (Frontline Systems, 1999) was employed. Implementation of the regional calibration approach for 30 sites with an average record length of 30.4 years (365 months), took approximately 30 min on a 200 MHz Pentium.

## Calibration results

Figure 2 compares the goodness of fit of the monthly streamflows generated by the “abcd” model using the traditional at-site calibration approach and the regional calibration approach introduced herein. Figure 2 uses three different statistics to represent the goodness of fit of the calibrated monthly flows to the observed flows: (a) the coefficient of determination  $R^2$ , (b) the coefficient of variation of the model residuals computed using  $C_v(\varepsilon) = \sigma_\varepsilon/\mu_Q$  and (c) the percentage bias. Since the model residuals  $\varepsilon_i$  in equation (6) should have mean zero,  $\sigma_\varepsilon/\mu_\varepsilon$  is undefined. Instead,  $C_v(\varepsilon) = \sigma_\varepsilon/\mu_Q$  is computed by dividing by the mean of the monthly flows so that  $C_v(\varepsilon)$  represents the standard deviation of the residuals as a fraction of the mean monthly streamflow. One observes from Fig. 2, that in terms of both goodness-of-fit statistics  $R^2$  and  $C_v(\varepsilon)$ , the at-site calibration approach is nearly always an improvement over the



**Fig. 2** Comparison of the traditional at-site calibration and the regional calibration approach introduced in this study using the goodness-of-fit statistics: (a) coefficient of determination  $R^2$ ; (b) coefficient of variation of model errors,  $C_v(\varepsilon) = \sigma_\varepsilon/\mu_Q$ ; and (c) percent bias.

regional calibration approach. This is to be expected because the objective function in the at-site calibration algorithm seeks to obtain the “best possible” fit at each site. Nevertheless the goodness-of-fit values corresponding to these two different calibration approaches are quite similar.

Figure 2(c) documents the percent bias computed using the formula:

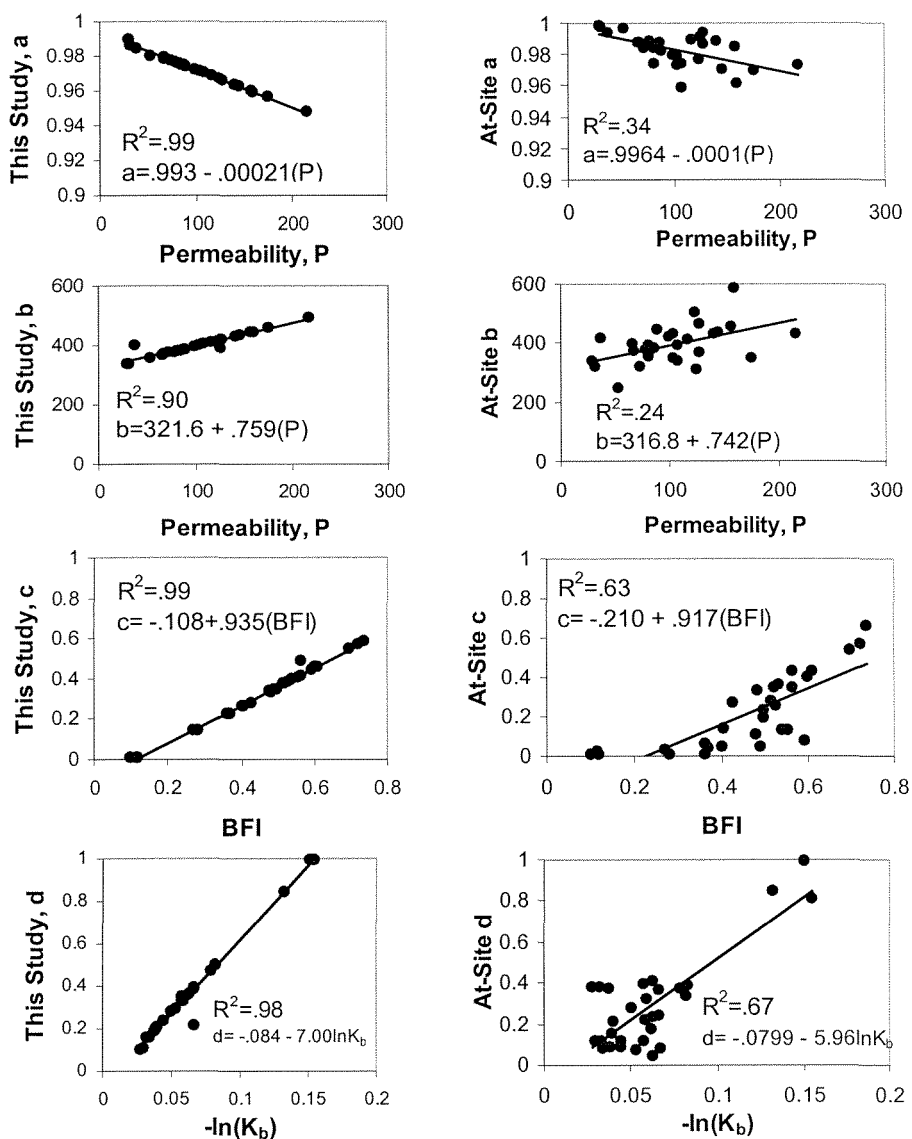
$$\%Bias = 100(\bar{\hat{Q}} - \bar{Q})/\bar{Q}$$

where  $\bar{\hat{Q}}$  represents the mean of the model generated flows and  $\bar{Q}$  represents the mean of the observed flows. Both calibration methods often result in bias because it is possible to obtain high values of  $R^2$  even for a biased model. Overall, the regional calibration approach led, on average, to unbiased models for the entire region, while the traditional at-site calibration approach resulted in upward bias. For individual sites, both approaches led to roughly the same variability in  $\%Bias$ .

Figure 3 illustrates the remarkably precise relationships between calibrated model parameters and watershed characteristics which result from using a regional calibration strategy. When an at-site calibration strategy is applied, the right-hand panels of Fig. 3 illustrate that the relationships between “abcd” model parameters and watershed characteristics are extremely weak. This result is consistent with the dozens of previous watershed model regionalization studies cited by Servat & Dezetter (1993), Abdulla & Lettenmaier (1997), Sefton & Howarth (1998), Xu & Singh (1998), Post *et al.* (1998), Xu (1999), Seibert (1999), Post & Jakeman (1999) and others. In the left-hand panels of Fig. 3 it is apparent that the regional calibration approach can produce extremely accurate regional regression relationships between watershed model parameters and watershed characteristics, while maintaining a goodness of fit between modelled and observed streamflows which is nearly as accurate as the best fit one can possibly achieve using an at-site algorithm. As is shown below, these nearly perfect regional regression relationships obtained using the new regional calibration approach, are misleading, because they do not result in improvements in the ability to calibrate a watershed model at an ungauged site!

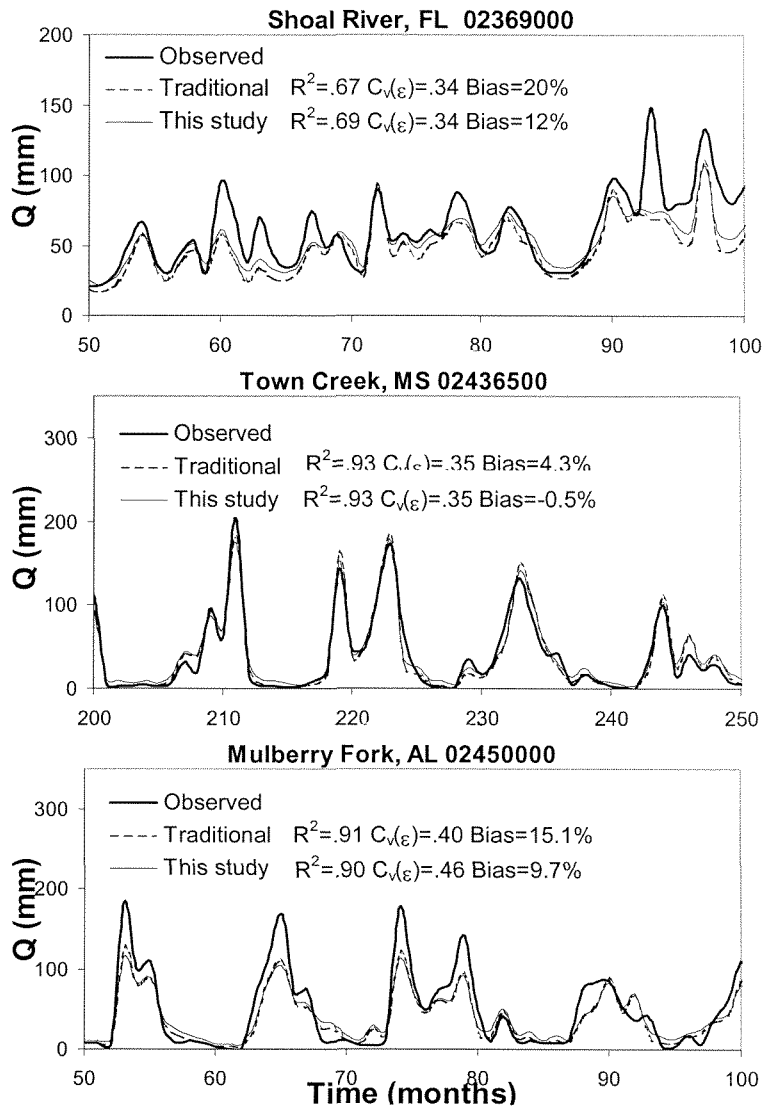
## MODEL VALIDATION

Research on hydrological watershed models has evolved considerably, along with the awareness that model structures and their associated model parameter sets are not unique, and infinite plausible mathematical representations exist. It is now generally understood that one can never validate a watershed model, rather, one can only invalidate it (Kirchner *et al.*, 1996). In an effort to invalidate the regional calibration approach introduced here, the methodology is evaluated using three basins which were not used to calibrate the model. The regional relationships between model parameters and watershed characteristics illustrated in Fig. 3 were used to estimate watershed model parameters and to generate monthly streamflows at three validation sites and the results are illustrated in Fig. 4. Here one observes that the traditional two-step regionalization approach produces nearly identical results to the regional calibration introduced in this study at all three validation sites. This result stems from the fact that



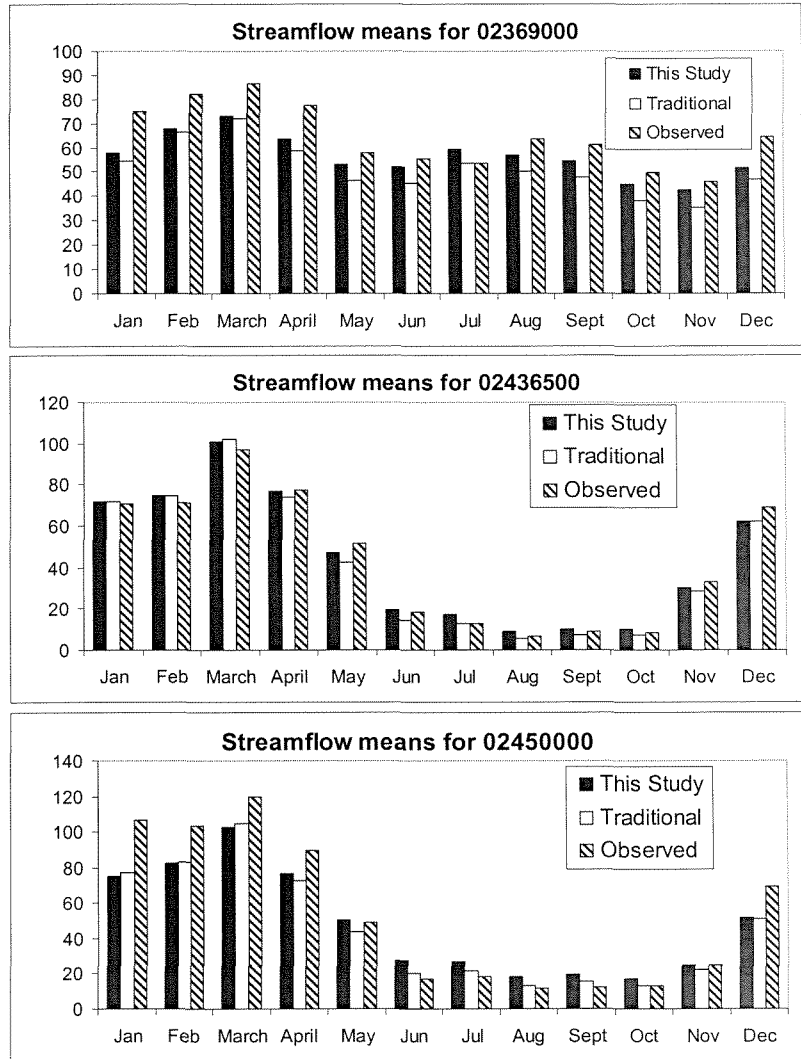
**Fig. 3** Relationships between “abcd” model parameters and watershed characteristics corresponding to at-site (right-hand panels) and regional calibration (left-hand panels) approaches.

the regional regression relationships between the model parameters and basin characteristics reported in Fig. 3 produce very similar regional relationships. So in spite of the fact that the traditional two-step regionalization approach leads to weak relationships between model parameters and basin characteristics, the relationships that do result are roughly equivalent to the tighter relationships produced by the regional calibration method.



**Fig. 4** Comparison of simulated and observed monthly streamflows corresponding to the traditional regional approach and the regional calibration approach introduced in this study for three validation sites.

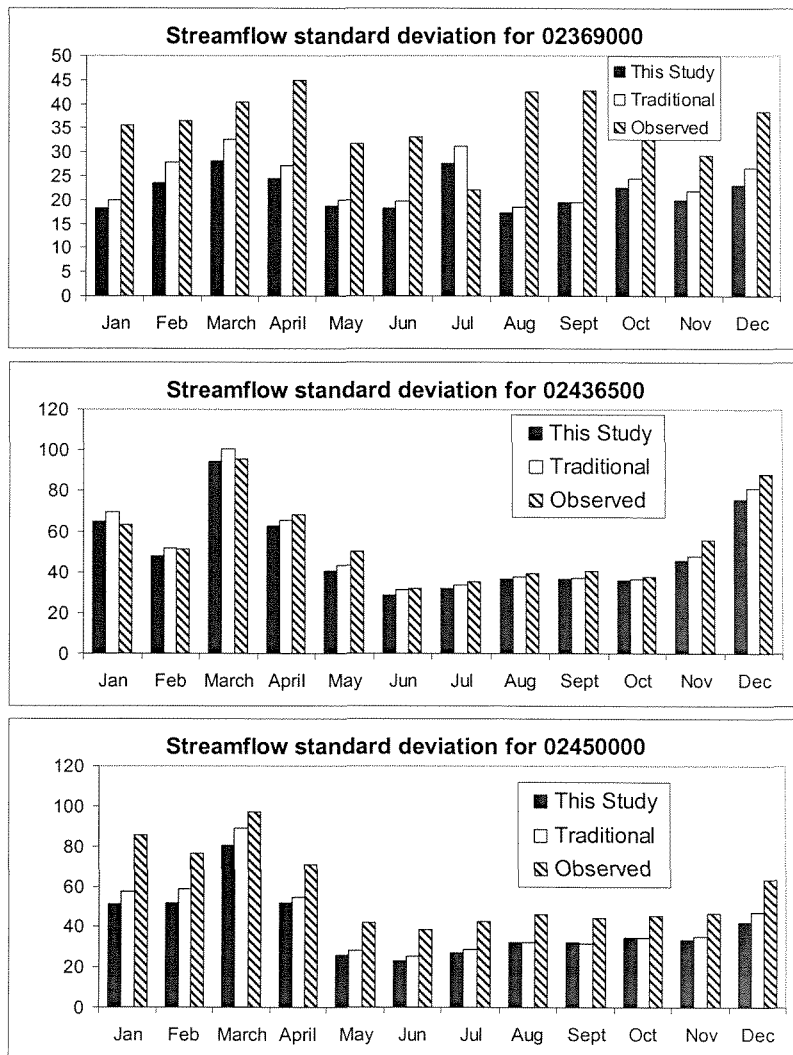
Figures 5 and 6 compare observed and modelled estimates of the mean and standard deviation, respectively, of the monthly streamflows at the three validation sites. One observes from Fig. 5 that both the traditional two-step regionalization approach and the regional calibration approach introduced herein reproduce the observed mean monthly streamflows with about the same accuracy at all three validation sites. As expected, both regionalization approaches tend to underestimate the standard deviation of the monthly streamflows as shown in Fig. 6. This is a general problem with all watershed models which can be proven as follows. Regardless of the model structure or temporal scale, streamflow can be expressed as:



**Fig. 5** Comparison of the observed and modelled mean monthly streamflows for the three validation sites.

$$Q_t = f(P_t, PE_t | \theta) + \varepsilon_t = \hat{Q}_t + \varepsilon_t$$

where  $f(P_t, PE_t | \theta)$  denotes the deterministic watershed model with inputs  $P_t$  and  $PE_t$ , model parameter set  $\theta$  and model error  $\varepsilon_t$  and  $\hat{Q}_t$  denotes modelled streamflow. When model error is independent of the model  $\text{Var}[\hat{Q}_t] = \text{Var}[Q_t] - \text{Var}[\varepsilon_t]$  so that in general,  $\text{Var}[\hat{Q}_t] < \text{Var}[Q_t]$  with the inequality becoming more important as model error increases. Therefore only a watershed model without error will be able to reproduce the standard deviation of the observed streamflows.



**Fig. 6** Comparison of the standard deviations of the observed and modelled monthly streamflows for the three validation sites.

## CONCLUSIONS

Given the increasingly widespread usage of watershed models for solving environmental problems, the regionalization of watershed models may be one of the most challenging and fundamental problems within the entire field of hydrology. Previous regionalization efforts have used a two-step approach where (a) the hydrological model is fitted to each site in a region and then (b) hydrological model parameters are related to watershed characteristics. Dozens of such regionalization attempts have been made by previous investigators for hourly, daily, and monthly watershed models, all producing mixed results. Invariably the relationships between watershed model

parameters and watershed characteristics are extremely weak, as was illustrated again in this study in the right-hand panel of Fig. 3. This study introduced a new approach to the regionalization of a hydrological model. The approach, termed regional calibration, attempts to calibrate the model at all sites in a region simultaneously, while concurrently attempting to achieve the best possible regional relationships among watershed model parameters and watershed characteristics. This approach led to remarkable improvements in the precision of the regional relationships between watershed model parameters and watershed characteristics when compared to the traditional two-step approach. However, the “remarkable” regional relationships corresponding to the regional calibration approach were later shown to be misleading, because validation experiments documented that both the traditional two-step approach and the regional calibration approach produce roughly equivalent streamflow simulations at three validation sites.

It is unclear on the basis of this initial experiment, whether the regional calibration methodology introduced in this study will ever lead to improvements in the ability to regionalize a watershed model over the traditional two-step approach. Future experiments with different hydrological models over a wide range of hydrological regimes may enable one to distinguish the conditions under which the regional calibration approach introduced here will (or will not) lead to improvements over the traditional two-step regionalization approach. The following remarks and conclusions should assist subsequent investigations which employ the regional calibration strategy introduced here:

- (a) It is well known that there exists a very large number of watershed model parameter sets which can produce physically realistic simulations. The regional calibration approach provides an attractive method for reducing the feasible subspace over which the watershed model calibration is performed. Kuczera (1997) documented that the addition of such constraints can lead to significant efficiencies in the overall watershed model parameter optimization problem. Even the regional calibration approach, with its inherent ability to reduce the feasible subspace of model parameters, results in a nonunique set of watershed model parameters. This apparent, yet ubiquitous, nonuniqueness of watershed model parameter sets will always confound our ability to estimate regional relationships between watershed model parameters and watershed characteristics.
- (b) The regional calibration approach can lead to dramatic improvements in the goodness-of-fit of regional relationships between watershed model parameters and basin characteristics. However, such remarkably good regional relationships imply significant underestimation of the uncertainty associated with resulting watershed model parameter estimates at ungauged sites. Future attempts to implement the regional calibration strategy need to properly account for the statistical properties of the model error and the multicollinearity among the watershed model parameters and the multicollinearity among the basin descriptors to properly account for the uncertainty associated with the resulting regional relationships.
- (c) In this preliminary evaluation of the regional calibration approach the multicollinearity among watershed model parameter estimates, as well as among the dependent variables in the regional regression relationships, was ignored. Tung *et al.* (1997) document the importance of accounting for the covariance structure of the watershed model parameters when fitting regional regression models. Future



research will hopefully combine the regional calibration idea introduced here with methods such as seemingly unrelated regression (Tung *et al.*, 1997) and generalized least squares regression (Kroll & Stedinger, 1998), which can account for the covariance structure of both the dependent and independent variables when fitting regional regression equations.

- (d) Previous regionalization studies have taken a different approach from that outlined in this study. Most previous studies have considered as their primary goal, estimation of watershed model parameters at ungauged sites. The results of such studies are rarely definitive because one never knows whether the goodness of fit of the regional relationships between watershed model parameters and basin characteristics could be improved by gathering better drainage basin information or by reformulating the structure of the regional relationships between watershed model parameters and basin characteristics. This study took a different approach. It set out only to develop regional relationships between watershed model parameters and basin characteristics under idealized conditions when streamflow records are available for estimating some of the basin characteristics. If one cannot solve this “data-rich” problem, one cannot hope to solve the “data-poor” (ungauged site) problem. It was found that even though the regional calibration method introduced here appears to offer significant potential for improving relationships between basin characteristics and watershed model parameters, it will not necessarily offer improvements in the ability to estimate model parameters at ungauged sites. If this experiment had been attempted for the “data-poor” (ungauged site) problem, this point could not have been proved. It is recommended that future researchers consider solving the “data-rich” problem described here, or a variant thereof, before attempting to solve the “data-poor” problem which exists at a purely ungauged site.
- (e) The regional calibration methodology introduced in this study could be extended to any problem involving the regionalization of a hydrological model. Other regional hydrological problems include regional flood flow and low flow frequency analysis and the regionalization of stochastic streamflow models.
- (f) Some of the most significant improvements in watershed modelling over the past few decades resulted from improvements in the ability to conceptualize and model hydrological processes. Ultimately, improvements in the ability to regionalize watershed models will only come after hydrologists begin to conceptualize and model regional physical relationships between watershed model parameters and watershed characteristics. As was clearly demonstrated by Wallis (1965), multivariate regression methods are unable to uncover basic physical laws. In other words, until hydrologists formulate the basic theoretical (physical) relationships between watershed model parameters and watershed characteristics, regionalization studies will continue to produce mixed results. Vogel & Kroll (1996) demonstrated this concept to the analogous problem of estimating regional hydrological models of low flow. They showed that improvements in regional models of low flow can be obtained by formulating spatial theoretical relationships among watershed model parameters and landscape attributes. One could provide citations to hundreds (possibly thousands) of different physically-based watershed simulation models. Interestingly, the authors are unaware of any studies which formulate physically-based regional hydrological relationships between watershed simulation model parameters and their associated landscape attributes. Given this

fact, it should come as no surprise that previous watershed model regionalization studies have met with limited success.

**Acknowledgements** Although the research described in this article has been funded in part by the US Environmental Protection Agency through grant no. R 824992-01-0 to Tufts University, it has not been subjected to the Agency's required peer and policy review and therefore does not necessarily reflect the views of the Agency and no endorsement should be inferred. The authors are indebted to Ian Wilson and Chris Daly for their assistance in the development of the monthly time series of precipitation, temperature and potential evapotranspiration. The authors are also indebted to Chuck Kroll for providing estimates of the baseflow recession constant and to Charlotte Bouvier for her translation of the abstract into French. The authors wish to thank Hoshin Gupta, Keith Beven, Zbigniew Kundzewicz and two anonymous reviewers for their comments which led to considerable improvements in the manuscript. A shortened version of this study was presented as a poster and oral presentation at the Workshop on the Regionalization of Parameters of Hydrological and Atmospheric Land Surface Models held at the general assembly of the International Union of Geodesy and Geophysics (IUGG) at the University of Birmingham, UK, on 27–28 July 1999.

## REFERENCES

- Abdulla, F. A. & Lettenmaier, D. P. (1997) Development of regional parameter estimation equations for a macroscale hydrologic model. *J. Hydrol.* **197**, 230–257.
- Allen, R. G. (1993) Evaluation of a temperature difference method for computing grass reference evapotranspiration. Report to Water Resources Development and Management Service Land and Water Development Division, FAO, Rome, Italy.
- Alley, W. M. (1984) On the treatment of evapotranspiration, soil moisture accounting, and aquifer recharge in monthly water balance models. *Wat. Resour. Res.* **20**, 1137–1149.
- Alley, W. M. (1985) Water balance models in one-month-ahead stream flow forecasting. *Wat. Resour. Res.* **21**(4), 597–606.
- Burn, D. H. & Boorman, D. B. (1992) Estimation of hydrological parameters at ungauged catchments. *J. Hydrol.* **143**, 429–454.
- Daly, C., Neilson, R. P. & Philips, D. L. (1992) A statistical-topographic model for mapping climatological precipitation over mountainous terrain. *J. Appl. Met.* **33**(2), 140–158.
- Daly, C., Taylor, G. & Gibson, W. (1997) The PRISM approach to mapping precipitation and temperature. In: *Proc. Tenth AMS Conf. on Applied Climatology* (Reno, 20–24 October 1997). Am. Met. Soc., Reno, Nevada, USA.
- Duan, Q., Sorooshian, S. & Gupta, V. (1992) Effective and efficient global optimization for conceptual rainfall-runoff models. *Wat. Resour. Res.* **28**(4), 1015–1031.
- Duffie, J. A. & Beckman, W. A. (1980) *Solar Engineering of Thermal Processes*, 1–109. Wiley, New York, USA.
- Fiering, M. B. (1967) *Streamflow Synthesis*. Harvard University Press, Cambridge, Massachusetts, USA.
- Frontline Systems (1999) *Premium Solver Plus Version 3.5 for Excel*. Frontline Systems Inc., PO Box 4288, Incline Village, Nevada, USA.
- Gustard, A., Bullock, A. & Dixon, J. M. (1992) Low flow estimation in the United Kingdom. Institute of Hydrology Report no. 108, Wallingford, UK.
- Hargreaves, G. H. & Samani, Z. A. (1982) Estimating potential evapotranspiration. *J. Irrig. Drain. Engng* **108**(3), 225–230.
- Institute of Hydrology (1980) Low flow studies. Report no. 1, Section 3.1, Institute of Hydrology, Wallingford, UK.
- Jensen, M. E., Burman, R. D. & Allen, R. G. (1990) *Evapotranspiration and Irrigation Water Requirements*. ASCE Manuals and Reports on Engineering Practice no. 70.
- Kirchner, J. W., Hooper, R. P., Kendall, C., Neal, C. & Leavesley, G. (1996) Testing and validating environmental models. *Sci. Total Environ.* **183**, 33–47.
- Kuczera, G. (1997) Efficient subspace probabilistic parameter optimization for catchment models. *Wat. Resour. Res.* **33**(1), 177–185.
- Kuczera, G. & Mroczkowski, M. (1998) Assessment of hydrologic parameter uncertainty and the worth of multiresponse data. *Wat. Resour. Res.* **34**(6), 1481–1489.

- Nathan, R. J. & McMahon, T. A. (1990) Evaluation of automated techniques for baseflow and recession analyses. *Wat. Resour. Res.* **26**(7), 1465–1473.
- Post, D. A. & Jakeman, A. J. (1999) Predicting the daily streamflow of ungauged catchments in S.E. Australia by regionalizing the parameters of a lumped conceptual rainfall–runoff model. *Ecol. Modelling* **123**, 91–104.
- Post, D. A., Jones, J. A. & Grant, G. E. (1998) An improved methodology for predicting the daily response of ungauged catchments. *Environ. Modelling & Software* **13**, 395–403.
- Rutledge, A. T. & Daniel, C. C., III (1994) Testing an automated method to estimate ground-water recharge from streamflow records. *Groundwater* **32**(2), 180–189.
- Sefton, C. E. M. & Howarth, S. M. (1998) Relationships between dynamic response characteristics and physical descriptors of catchments in England and Wales. *J. Hydrol.* **211**, 1–16.
- Seibert, J. (1999) Regionalization of parameters for a conceptual rainfall–runoff model. *Agric. For. Met.* **98–99**, 279–293.
- Servat, E. & Dezetter, A. (1993) Rainfall–runoff modeling and water resources assessment in northwestern Ivory Coast, Tentative extension to ungauged catchments. *J. Hydrol.* **148**, 231–248.
- Shuttleworth, W. J. (1993) Evaporation. In: *Handbook of Hydrology* (ed. by D. R. Maidment), 4.18. McGraw-Hill, New York, USA.
- Slack, J. R., Lumb, A. M. & Landwehr, J. M. (1993) Hydroclimatic data network (HCDN): as US Geological Survey streamflow data set for the United States for the study of climate variation, 1874–1988. Water-Resource Investigations Report no. 93–4076, US Geological Survey, Washington, DC, USA.
- Taylor, G. H., Daly, C., Gibson, W. P. & Sibul-Weisberg, J. (1997) Digital and map products produced using PRISM. In: *Proc. Tenth AMS Conf. on Applied Climatology* (Reno, 20–24 October 1997). Am. Met. Soc., Reno, Nevada, USA.
- Thomas, H. A. (1981) Improved methods for national water assessment. Report, Contract WR 15249270, US Water Resources Council, Washington, DC, USA.
- Thomas, H. A., Marin, C. M., Brown, M. J. & Fiering, M. B. (1983) Methodology for water resource assessment. Report NTIS 84–124163, to US Geological Survey, National. Tech. Info. Serv., Springfield, Virginia, USA.
- Tung, Y.-K., Yeh, K.-C. & Yang, J.-C. (1997) Regionalization of unit hydrograph parameters, 1. comparison of regression analysis techniques. *Stochast. Hydrol. & Hydraul.* **11**, 145–171.
- Vandewiele, G. L., Xu, C.-Y. & Ni-Lar-Win (1992) Methodology and comparative study of monthly water balance models in Belgium, China and Burma. *J. Hydrol.* **134**, 315–347.
- Vandewiele, G. L. & Elias, A. (1995) Monthly water balance of ungauged catchments obtained by geographical regionalization. *J. Hydrol.* **170**, 277–291.
- Vandewiele, G. L. & Ni-Lar-Win (1998) Monthly water balance models for 55 basins in 10 countries. *Hydrol. Sci. J.* **43**(5), 687–699.
- Vogel, R. M. & Kroll, C. N. (1992) Regional geohydrologic-geomorphic relationships for estimation of low-flow statistics. *Wat. Resour. Res.* **28**(9), 2451–2458.
- Vogel, R. M. & Kroll, C. N. (1996) Estimation of baseflow recession constants. *Wat. Resour. Manage.* **10**, 303–320.
- Wallis, J. R. (1965) Multivariate statistical methods in hydrology—a comparison using data of known functional relationship. *Wat. Resour. Res.* **1**(4), 447–461.
- Wolock, D. M. (1997) STATSGO soil characteristics for the conterminous US. USGS Open-File Report 97–656.
- Xu, C.-Y. (1999) Estimation of parameters of a conceptual water balance model for ungauged catchments. *Wat. Resour. Manage.* **13**, 353–368.
- Xu, C.-Y. & Singh, V. P. (1998) A review of monthly water balance models for water resource investigations. *Wat. Resour. Manage.* **12**, 31–50.

Received 2 September 1999; accepted 5 July 2000

